Démonstration élémentaire d'un principe de la méthode des moindres carrés

DÉMONSTRATION ÉLÉMENTAIRE

d'un principe de la méthode des moindres carrés

PAR E. LE GRAND ROY, PROF.

On sait que, pour tirer de m équations linéaires à i inconnues, dans le cas de $m > i$, les valeurs les plus probables de ces inconnues, on emploie la méthode dite « des moindres carrés », dont il est nécessaire de rappeler brièvement le principe.

$$\text{Soient} \left\{ \begin{array}{l} ax \;\; + by \;\; + cz \;\; \ldots = n \\ a'x \;\; + b'y \;\; + c'z \;\; \ldots = n' \\ a''x + b''y + c''z \ldots = n'' \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \end{array} \right\} \quad (1)$$

les m équations à résoudre dans lesquelles n, n', n''... sont des nombres obtenus par observation. On les remplace par i équations nouvelles, qui s'obtiennent en multipliant chaque équation par le coefficient d'une même inconnue pris dans cette équation, et additionnant les résultats obtenus. La première s'obtiendra donc en multipliant chaque équation par le coefficient de x qu'elle renferme; la seconde, en multipliant chaque équation par le coefficient de y, et ainsi de suite. Si donc on pose, pour abréger :

$$(2)\left\{\begin{array}{ll}a^2+a'^2+a''^2+\ldots=(aa) & ba+b'a'+b''a''+\ldots=(ba)=(ab)\\ ab+a'b'+a''b''+\ldots=(ab) & b^2+b'^2+b''^2+\ldots=(bb)\\ ac+a'c'+a''c''+\ldots=(ac) & bc+b'c'+b''c''+\ldots=(bc)\\ \ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\\ an+a'n'+a''n''+\ldots=(an) & bn+b'n'+b''n''+\ldots=(bn)\end{array}\right.$$

et ainsi de suite, on obtient comme équations finales, en nombre égal à celui des inconnues :

$$(3)\left\{\begin{array}{l}(aa)x+(ab)y+(ac)z+\ldots=(an)\\ (ba)x+(bb)y+(bc)z+\ldots=(bn)\\ (ca)x+(cb)y+(cc)z+\ldots=(cn)\\ \ldots\ldots\ldots\ldots\ldots\end{array}\right.$$

Les valeurs tirées de ces équations ne vérifient pas rigoureusement les équations (1) : tandis qu'on devrait avoir, si les valeurs observées et les valeurs des inconnues étaient rigoureusement exactes :

$$\left\{\begin{array}{l}ax+by+cz\ldots-n=0\\ a'x+b'y+c'z\ldots-n'=0\\ a''x+b''y+c''z\ldots-n''=0\\ \ldots\ldots\ldots\ldots\ldots\end{array}\right.$$

la substitution dans les équations (1), ainsi modifiées, des valeurs des inconnues tirées des équations (3) donne :

$$(4)\left\{\begin{array}{l}ax+by+cz\ldots-n=\mathrm{E}\\ a'x+b'y+c'z\ldots-n'=\mathrm{E}'\\ a''x+b''y+c''z\ldots-n''=\mathrm{E}''\\ \ldots\ldots\ldots\ldots\ldots\end{array}\right.$$

les *résidus* E, E′, E″... étant d'autant plus petits que les observations sont plus exactes. On prend pour mesure de leur exactitude leur *erreur moyenne*, qui se calcule par la formule

$$E_1 = \pm \sqrt{\frac{E^2 + E'^2 + E''^2 + \ldots}{m - i}}$$

ou, en posant

$$E^2 + E'^2 + E''^2 + \ldots = (EE) \qquad (5)\ E_1 = \pm \sqrt{\frac{(EE)}{m - i}}$$

Pour le calcul de l'erreur moyenne de chaque inconnue, on applique la règle suivante : dans les équations (3), on accentue les lettres $x, y, z \ldots$, et on remplace les seconds membres par 1 dans l'une des équations, par 0 dans toutes les autres. On obtient ainsi les systèmes d'équations :

$$(6)\ \left\{\begin{array}{l} (aa)x' + (ab)y' + (ac)z' \ldots = 1 \\ (ba)x' + (bb)y' + (bc)z' \ldots = 0 \\ (ca)x' + (cb)y' + (cc)z' \ldots = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

$$(6\ \text{bis})\ \left\{\begin{array}{l} (aa)x' + (ab)y' + (ac)z' \ldots = 0 \\ (ba)x' + (bb)y' + (bc)z' \ldots = 1 \\ (ca)x' + (cb)y' + (cc)z' \ldots = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

$$(6\ \text{ter})\ \left\{\begin{array}{l} (aa)x' + (ab)y' + (ac)z' \ldots = 0 \\ (ba)x' + (bb)y' + (bc)z' \ldots = 0 \\ (ca)x' + (cb)y' + (cc)z' \ldots = 1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

en nombre égal à celui des inconnues.

Si alors on tire de (6) la valeur de x', de (6 bis) celle de y', de (6 ter) celle de $z' \ldots$, et qu'on appelle E_x, E_y, $E_z \ldots$ les erreurs moyennes des inconnues, on les obtient par les formules

$$(7)\ E_x = \pm E_1 \sqrt{x'} \qquad E_y = \pm E_1 \sqrt{y'} \qquad E_z = \pm E_1 \sqrt{z'}$$

C'est de ces formules que nous avons cherché une démonstration nouvelle, celles que donnent les différents ouvrages qui traitent de cette question nous ayant paru manquer, soit de clarté, soit de généralité.

Résolvons par rapport à x les équations (3), où nous introduisons une 4me inconnue u. On a ainsi les équations

$$(3 \text{ bis}) \begin{cases} (aa)x + (ab)y + (ac)z + (ad)u \ldots = (an) \\ (ba)x + (bb)y + (bc)z + (bd)u \ldots = (bn) \\ (ca)x + (cb)y + (cc)z + (cd)u \ldots = (cn) \\ (da)x + (db)y + (dc)z + (dd)u \ldots = (dn) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

d'où l'on tire

$$(8)\ x = \frac{\begin{vmatrix} (an) & (ab) & (ac) & (ab). \\ (bn) & (bb) & (bc) & (bd). \\ (cn) & (cb) & (cc) & (cb). \\ (dn) & (db) & (dc) & (dd). \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}}{\begin{vmatrix} (aa) & (ab) & (ac) & (ad). \\ (ba) & (bb) & (bc) & (bd). \\ (ca) & (cb) & (cc) & (cd). \\ (da) & (db) & (dc) & (dd). \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}} = \frac{\begin{vmatrix} (an) & (ab) & (ac) & (ad). \\ (bn) & (bb) & (bc) & (bd). \\ (cn) & (cb) & (cc) & (cd). \\ (dn) & (db) & (dc) & (dd). \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}}{\mathrm{D}}$$

D désignant le déterminant du système (3 bis).

Désignons par $\Delta_1, \Delta_2, \Delta_3 \ldots$ les déterminants mineurs relatifs aux termes de la 1re colonne. On a alors, en développant :

$$\mathrm{D}x = (an)\Delta_1 - (bn)\Delta_2 + (cn)\Delta_3 - (dn)\Delta_4 + \ldots$$

ou, en vertu des formules (2) :

$$Dx = (an + a'n' + a''n'' + \ldots)\,\Delta_1 - (bn + b'n' + b''n''\ldots)\Delta_2 + \ldots = (a\Delta_1 - b\Delta_2 + c\Delta_3 - d\Delta_4\ldots)n + (a'\Delta_1 - b'\Delta_2 + c'\Delta_3\ldots)n' + (a''\Delta_1 - b''\Delta_2 + c''\Delta_3\ldots)n'' + \ldots$$

Différentions les deux membres et remplaçons les différentielles de x, n, n'... par les erreurs E_x, E, E′... On obtient :

$$DE_x = (a\Delta_1 - b\Delta_2 + c\Delta_3 - d\Delta_4\ldots)E + (a'\Delta_1 - b'\Delta_2 + c'\Delta_3 - d'\Delta_4\ldots)E' + (a''\Delta_1 - b''\Delta_2 + c''\Delta_3 - d''\Delta_4\ldots)E'' + \ldots$$

Elevons les deux membres au carré, et remarquons que, les erreurs accidentelles étant indifféremment positives ou négatives, à chaque double produit tel que E E′ en correspondra un autre égal et de signe contraire. On peut donc négliger les termes qui renferment ces doubles produits, qui se détruisent sensiblement 2 à 2, et on obtient :

$$D^2E_x^2 = (a\Delta_1 - b\Delta_2 + c\Delta_3 - d\Delta_4\ldots)^2E^2 + (a'\Delta_1 - b'\Delta_2 + c'\Delta_3 - d'\Delta_4\ldots)^2E'^2 + (a''\Delta_1 - b''\Delta_2 + c''\Delta_3 - c''\Delta_4\ldots)^2E''^2 + \ldots$$

Remplaçons chacune des erreurs E, E′, E″... par l'erreur moyenne E_1. On a alors :

$$D^2E_x^2 = E_1^2\left[(a\Delta_1 - b\Delta_2 + c\Delta_3\ldots)^2 + (a'\Delta_1 - b'\Delta_2 + c'\Delta_3\ldots)^2 + (a''\Delta_1 - b''\Delta_2 + c''\Delta_3\ldots)^2 + \ldots\right]$$

ou, en développant :

$$D^2E_x^2 = E_1^2\left[(a^2 + a'^2 + a''^2\ldots)\Delta_1^2 - 2(ab + a'b' + a''b'' + \ldots)\Delta_1\Delta_2 + b^2 + b'^2 + b''^2\ldots)\Delta_2^2 + 2(ac + a'c' + a''c''\ldots)\Delta_1\Delta_3 - 2(bc + b'c' + b''c''\ldots)\Delta_2\Delta_3 + (c^2 + c'^2 + c''^2 + \ldots)\Delta_3^2 - 2(ad + a'd' + a''d''\ldots)\Delta_1\Delta_4 + 2(bd + b'd' + b''d''\ldots)\Delta_2\Delta_4 - 2(cd + c'd' + c''d''\ldots)\Delta_3\Delta_4 + (d^2 + d'^2 + d''^2\ldots)\Delta_4^2\ldots\right]$$

ou, en introduisant les notations abrégées des formules (2),

$$D^2 E_x^2 = E_1^2 \left[(aa)\Delta_1^2 - 2(ab)\Delta_1\Delta_2 + (bb)\Delta_2^2 + 2(ac)\Delta_1\Delta_3 - 2(bc)\Delta_2\Delta_3 + (cc)\Delta_3^2 - 2(ad)\Delta_1\Delta_4 + 2(bd)\Delta_2\Delta_4 - 2(cd)\Delta_3\Delta_4 + (dd)\Delta_4^2 + \ldots \right]$$

ce qui peut s'écrire aussi

$$D^2 E_x^2 = E_1^2 \left[(aa)\Delta_1^2 - (ab)\Delta_1\Delta_2 - (ab)\Delta_1\Delta_2 + (bb)\Delta_2^2 + (ac)\Delta_1\Delta_3 + (ac)\Delta_1\Delta_3 - (bc)\Delta_2\Delta_3 - (bc)\Delta_2\Delta_3 + (cc)\Delta_3^2 - (ad)\Delta_1\Delta_4 - (ad)\Delta_1\Delta_4 + (bd)\Delta_2\Delta_4 + (bd)\Delta_2\Delta_4 - (cd)\Delta_3\Delta_4 - (cd)\Delta_3\Delta_4 + (dd)\Delta_4^2 \ldots \right]$$

ou encore

$$D^2 E_x^2 = E_1^2 \Big\{ \left[(aa)\Delta_1 - (ab)\Delta_2 + (ac)\Delta_3 - (ad)\Delta_4 + \ldots \right]\Delta_1 - \left[(ab)\Delta_1 - (bb)\Delta_2 + (bc)\Delta_3 - (bd)\Delta_4 + \ldots \right]\Delta_2 + \left[(ac)\Delta_1 - (bc)\Delta_2 + (cc)\Delta_3 - (cd)\Delta_4 \ldots \right]\Delta_3 - \left[(ad)\Delta_1 - (bd)\Delta_2 + (cd)\Delta_3 - (dd)\Delta_4 \ldots \right]\Delta_4 + \ldots \Big\}$$

En se reportant à la formule (8), on voit que

$$D = \begin{vmatrix} (aa) & (ab) & (ac) & (ad) & . \\ (ba) & (bb) & (bc) & (bd) & . \\ (ca) & (cb) & (cc) & (cd) & . \\ (da) & (db) & (dc) & (dd) & . \\ . & . & . & . & . \end{vmatrix} = (aa)\Delta_1 - (ba)\Delta_2 + (ca)\Delta_3 - (da)\Delta_4 \ldots$$

Dans le second membre de l'équation précédente, le cœfficient de Δ_1 est donc égal à D. Ceux des quantités Δ_2, Δ_3, Δ_4... sont les sommes algébriques des produits des termes d'une des colonnes de D par les déterminants mineurs relatifs aux termes d'une autre colonne, et la théorie des déterminants montre que ces sommes sont nulles.

Cette équation se réduit donc à

$$D^2 E_x^2 = E_1^2 . D \Delta_1$$

d'où

$$E_x = \pm E_1 \sqrt{\frac{\Delta_1}{D}} \quad (9)$$

Reprenons maintenant la formule (8), et développons le numérateur suivant les termes de la première colonne.

On obtient :

$$x = \frac{(an)\Delta_1 - (bn)\Delta_2 + (cn)\Delta_3 - (dn)\Delta_4 + \ldots}{D}$$

(an), (bn), (cn) ... étant les 2ds membres des équations (3 bis). Mais, d'après les formules (7), on doit avoir :

$$E_x = \pm E_1 \sqrt{x'}$$

x' étant tiré du système

$$(10) \begin{cases} (aa)x' + (ab)y' + (ac)z' + (ad)u' \ldots = 1 \\ (ba)x' + (bb)y' + (bc)z' + (bd)u' \ldots = 0 \\ (ca)x' + (cb)y' + (cc)z' + (cd)u' \ldots = 0 \\ (da)x' + (db)y' + (dc)z' + (dd)u' \ldots = 0 \end{cases}$$

Il faut donc, si la règle est juste, que $x' = \frac{\Delta_1}{D}$.

Pour le vérifier, remarquons que ces équations ne diffèrent des équations (3 bis) que par leurs 2ds membres.

Si donc nous reprenons la formule

$$x = \frac{(an)\,\Delta_1 - (bn)\,\Delta_2 + (cn)\,\Delta_3 - (dn)\,\Delta_4 + \ldots}{D}$$

il suffit, pour qu'elle donne la valeur de x', d'y remplacer (an), (bn), $(cn)\ldots$, qui sont les 2ds membres des équations (3 bis), par 1, 0, 0... qui sont les 2ds membres des équations (10), et on obtient alors

$$x' = \frac{\Delta_1}{D}$$

On peut donc, dans la formule (9), remplacer $\frac{\Delta_1}{D}$ par x', ce qui donne $E_x = \pm E_1\sqrt{x'}$, c'est-à-dire précisément la 1re des formules (7). Les autres se justifieraient de la même manière.